May 2012

Physics Performance Data Mining

Andrew Richard Johns
Worcester Polytechnic Institute

Physics Performance Data Mining

An Interactive Qualifying Project Report submitted to the Faculty of the

Worcester Polytechnic Institute

In partial fulfillment of the requirements for the Degree of Bachelor of Science

By Andrew Johns

Advisor: Germano Iannacchione

Abstract

Physics is one of the most important fields of study in the world. By identifying correlations between students and their knowledge of physics we can greater understand how to keep advancing in physics as a society. Studying data is a very practical way to understand humans in relation to physics. Students that succeed in math will also succeed in physics. Understanding that correlation can help to improve the knowledge and learning of physics across the world.

Acknowledgements

I would like to take this time to thank my advisor Professor Germano Iannacchione, whom without his teaching and guidance this project would not have been completed. His love of physics inspired me to do this project to help improve the quality of physics here at Worcester Polytechnic Institute. His humor and ability to understand and interact with students made this project extremely fun and educational. I sincerely hope this project is the beginning of long standing study of how to better understand the learning of physics.

Table of contents

List of Figures

List of Tables

Executive Summary

This report is being written to provide a preliminary study of physics data here at Worcester Polytechnic Institute. It will include an introduction and background to the study of correlations of students and their success in physics. It will then illustrate the methodology of the data testing including what components were considering when testing the data. The results will be published with detailed figures and tables that show the significance of the testing. The report will finish with conclusions and recommendations that can be drawn from the data testing.

Introduction

The science of studying matter and energy, also known as physics, has been around for a very long time. Without the works of famous physicists such as Sir Isaac Newton, Albert Einstein, and Benjamin Franklin the world of physics may be a completely different place. Physics is one of the most important fields of study that there is today. Without physics advances in new technologies would be nearly impossible. Televisions, computers, appliances, and nuclear weapons are just some of the major advances in society due to physics. These reasons and many more have been my motivation to pursue this physics based interactive qualifying project. Physics is one of the core classes here at Worcester Polytechnic Institute and for good reason, almost every upper level class I have taken as an engineer has involved some physics related material. However, because physics is such a major part of school and life it is a very difficult subject to fully grasp. Physicists are still experimenting and discovering new concepts daily and therefore the study of physics is always changing. Many students here at Worcester Polytechnic Institute take advanced physics classes in high school only to find that there is so much more to be learned. Through this interactive qualifying project I will dissect physics data in order to look for correlations that may help to improve the quality of physics courses being taught here at Worcester Polytechnic Institute.

Background

There have been multiple studies done on the relationships between learning performances and humans and their characteristics. One specific relationship question has always been asked and has been studied by many scientists; are males smarter than females? Many say that there is no intellectual difference between genders however many studies show differently. Psychologist J. Philippe Rushton firmly believes through his studies that males are smarter than females. One of Rushton's experiments involved studying SAT scores that lead to an interesting conclusion. He discovered that males surpassed females by an average of 3.6 IQ points (Bryner April 2012). Still many believe that this is not the case and that there is evidence of females being smarter than males. Studies will continue to grow and continue to unveil interesting new discoveries about this topic. This project is focused on a college subject therefore there may be some other interesting characteristics to take into account. Along with gender some of the other characteristics I found that might be interesting to look at were class year, college major, and similar subject performance. Another interesting study that this project only slightly looked into that would be very intriguing to look into, would be the relationship between the performances of all students in relation to the professor. After finding some information on grades relating to professor you could look at teaching style along with testing techniques among many other variables. There are many ways to look at data and come up with hypothesis for why certain things came out the way they did. In this paper I will show my results and talk about how and why I did the testing that I did.

Methodology

To start the data mining process I needed the data that I could not access by myself. I had to contact the registrar's office asking permission to gain access to the data for an interactive qualifying project. I asked for any and as much data as possible because the more data points to compare the better the results. Pam Theodore of the Provost's Office worked with me to attain data including physics courses taken and grade received, calculus course taken concurrently with physics and the grade received, gender, graduating year, and major. Since I am the only student working on the project much of the data could not be tested as thoroughly as I would have liked. However I firmly believe my finding will be of great value to the physics department as a guideline to improving physics based courses. This project is in no way shape or form a human nature or behavior project. The sole purpose of this project is education and the improvement of education. I received no record of the names or any private information of any of the students involved in this study. I have strictly focused on major, gender, class year, and calculus vs. physics data and have found some interesting results. The data set given to me by the Provost's office includes any full-time undergraduate student who has taken an undergraduate physics course since the fall of 2006.

When looking at gender I took all of the physics grades and sorted them by female versus male. I then sorted those two groups by grade received in physics (A, B, C, and NR). The results are in percentages of A's among females versus A's among males and so on. When looking at major I grouped all of the majors here at WPI into seven main groups. The groups are Mathematics, Science, Computer, Engineering, Other, Humanities and Arts, and Business. Those groups were then studied by grade received in physics versus grade received in physics against a different major. Another subject I thought that might be interesting to look at was class year. I

divided all students into two main groups, freshman and all upperclassmen. Those two groups were studied by grade received in physics also. The last characteristic I studied was calculus grades versus physics grades. In this study I took basic physic course grades and compared them to basic calculus course grades that they were taking at the same time. I then looked at the comparison of grades in each class against one another.

Results and Discussion

Grade Trend

Table 1 and Figure 1 show the relationship from year to year of all physics class performances from all students. The data stretches from the 2006-2007 school year to the beginning of the 2011-2012 school year. Every E term physics class data is hard to include when looking at final results. This is because summer classes are much easier and more one on one because not many students take summer classes. Overall there was no distinct relationship from year to year or even term to term that would be worthy of a discussion of what happened. The only tiny bit of evidence that I found somewhat interesting was that every d term the amount of NR's seemed to increase slightly. This could be complete coincidence or it could be due to the fact that at the end of each year the amount of upper level physics courses being taken is much higher. It could also have something to do with the fact that students seem to slack off at the end of the year. Besides the analysis of those minor occurrences this data was not studied on a smaller scale. In a larger study I believe it would be interesting to know what professor taught during each term and see if different professors have major impacts with good or bad grades. In my experience here at Worcester Polytechnic Institute the professor has had a large impact on not only my final grade but the amount I attended class, pay attention, do homework, and how enthusiastic I am about the material. The appendix has four graphs under the grade trend section the show the grade trend for separate physics courses. PH 1110 and PH 1111 show similar grade trend results to that of the overall physics grade trend while PH 2201 shows a higher trend in C's and NR's probably due to the fact that it is a harder course.

Count of SHVGPAI_GRADE	Term				
Row Labels	**NR**	**C**	**B**	**A**	**Grand Total**
2007A	15.82%	22.90%	34.85%	26.43%	100.00%
2007B	13.40%	21.69%	32.45%	32.45%	100.00%
2007C	15.45%	19.10%	32.30%	33.15%	100.00%
2007D	15.53%	22.89%	32.63%	28.95%	100.00%
2007E	5.56%	27.78%	50.00%	16.67%	100.00%
2008A	11.02%	26.58%	38.25%	24.15%	100.00%
2008B	12.14%	26.81%	29.68%	31.37%	100.00%
2008C	6.96%	21.45%	40.95%	30.64%	100.00%
2008D	17.90%	23.63%	28.16%	30.31%	100.00%
2008E	6.90%	20.69%	37.93%	34.48%	100.00%
2009A	8.53%	18.76%	40.31%	32.40%	100.00%
2009B	9.15%	22.24%	33.12%	35.49%	100.00%
2009C	10.34%	19.64%	32.30%	37.73%	100.00%
2009D	13.33%	26.67%	30.00%	30.00%	100.00%
2009E	16.67%	20.83%	33.33%	29.17%	100.00%
2010A	13.09%	24.41%	33.53%	28.97%	100.00%
2010B	11.36%	21.10%	33.44%	34.09%	100.00%
2010C	7.77%	19.66%	36.89%	35.68%	100.00%
2010D	17.14%	23.52%	30.77%	28.57%	100.00%

2010E2	0.00%	16.67%	45.83%	37.50%	100.00%
2011A	10.67%	17.84%	40.35%	31.14%	100.00%
2011B	7.80%	22.44%	38.37%	31.38%	100.00%
2011C	8.83%	21.00%	39.14%	31.03%	100.00%
2011D	10.07%	27.17%	33.26%	29.51%	100.00%
2011E2	10.71%	28.57%	35.71%	25.00%	100.00%
2012A	11.00%	22.70%	42.92%	23.38%	100.00%

Table 1: Grade Trend from 2007-2012

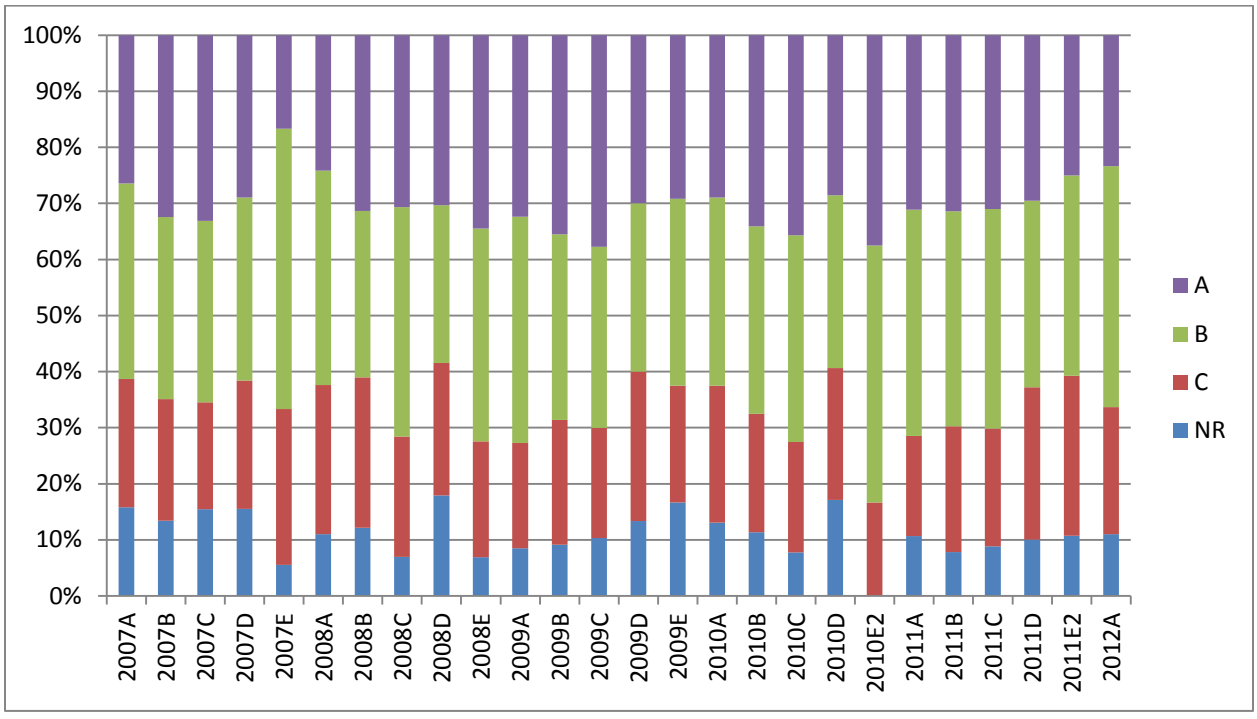


Figure 1: Grade Trend from 2007-2012

Gender

Gender is a very interesting topic when looking into comparisons of intelligence. Figure 2 and Table 2 show the percentages of each gender that received a specific grade. The gender graph is very interesting to me because of what the data shows. There is about a 2% higher failure rate among males than females. Also there is about a 2% increase in the amount of females who receive B's than males who receive B's. These results, although seem very low percentages, are a very large amount of the population considering the amount of data points that were used. However the results may not be concrete evidence that females are smarter or better at physics than males. There are a range of variables that have to be looked into, for example Worcester Polytechnic Institute is mostly males therefore the percentage of males and females who receive A's may be about the same however 30% of the male population is a lot larger number of students than 30% of the female population. Also even if the ratio to males and females were even, a 2% difference is not a large enough to make any sizeable conclusions. It is however very interesting to consider. If there was a way to take random samples from terms of the same number of students from each gender and then compare that data the results may look about the same or could be completely different and I believe that would be something very interesting to look into. However I do believe the results from my testing are rather interesting considering the fact that common perception is that males are smarter than females. I do not believe this to be the case, and even though the numbers of each gender aren't the same I still think the results show that there is no major intellectual difference between males and females. The appendix further breaks down the gender trend by upper level and lower level physics courses. The lower level physics course grades are roughly about the same for both males and

females however, the upper level physics courses show a distinct change. Females have a 10% higher rate in the amount of A's received in these courses. This is quite a large percentage and would be something that the professors here at Worcester Polytechnic Institute might want to look into.

Count of SHVGPAI_GRADE	**Column Labels**				
Row Labels	**NR**	**C**	**B**	**A**	**Grand Total**
F	9.75%	23.38%	37.40%	29.48%	100.00%
M	11.23%	23.28%	35.82%	29.67%	100.00%
Grand Total	**10.83%**	**23.30%**	**36.24%**	**29.62%**	**100.00%**

Table 2: Percentages of Grades by Gender

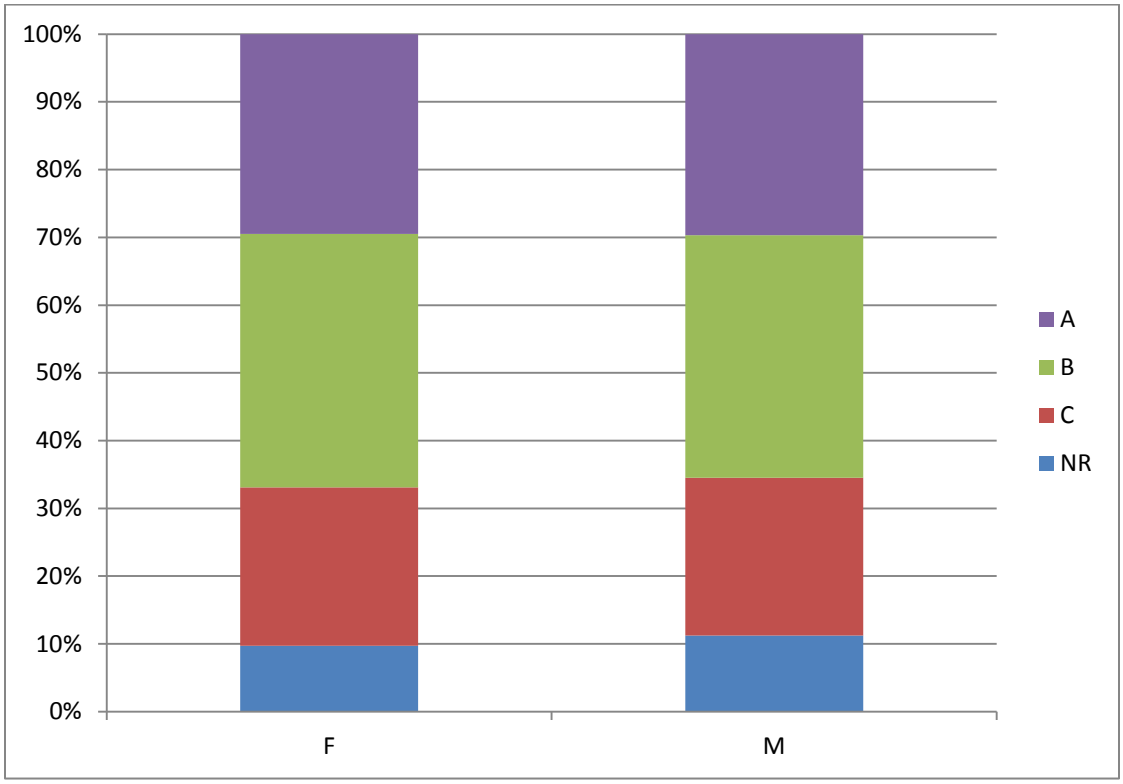


Figure 2: Male Performance versus Female Performance

Major

There are many subjects for students to major in when attending college. Worcester Polytechnic Institute has a very respectable reputation as being one of the best engineering schools in the nation. There are many different types of engineering to major in but most of them are strongly rooted in science and math based backgrounds. Worcester Polytechnic Institute, although not as well advertised, has a wide variety of other majors to choose from such as business and humanities. This study took a look into the performance differences between students with majors that are closely related to physics and those with majors who are not closely related to physics. Table 3 and Figure 3 show the number of students in each of the seven major categories and the grades they received. Looking at the data from a college major point of few some easily perceived results were found. Students with majors closely relating in the field of physics, such as science and mathematics, did much better than the students in the fields of humanities and business. However there were much more students majoring in engineering, computers, and science than in the other majors which is a variable that must be accounted for. Therefore the higher percentage of students in engineering and the lower percentage of students in business could alter the data. Without taking into account the number of students in each major, Figure 3 shows a very interesting trend especially in the amount of A's. The Math and Science majors excelled in physics while the humanities & arts and business majors struggled. The appendix for this section shows a histogram for upper level physics courses only, and the results are quite disturbing. Humanities and arts, business, other, and undeclared majors have little to almost no success in these classes. This is something that needs attention from the physics faculty, these students need a fair chance at success just like all the engineering, math and physics majors.

Count of SHVGPAI_GRADE	Column Labels				
Row Labels	**NR**	**C**	**B**	**A**	**Grand Total**
MAT	13	37	59	73	182
SCI	159	276	463	450	1348
COMP	112	153	228	243	736
ENG	675	1602	2494	1912	6683
OTH	39	74	105	74	292
HUSS	2	12	18	9	41
BUS	12	23	19	6	60
Grand Total	**1012**	**2177**	**3386**	**2767**	**9342**

Table 3: Grades by College Major

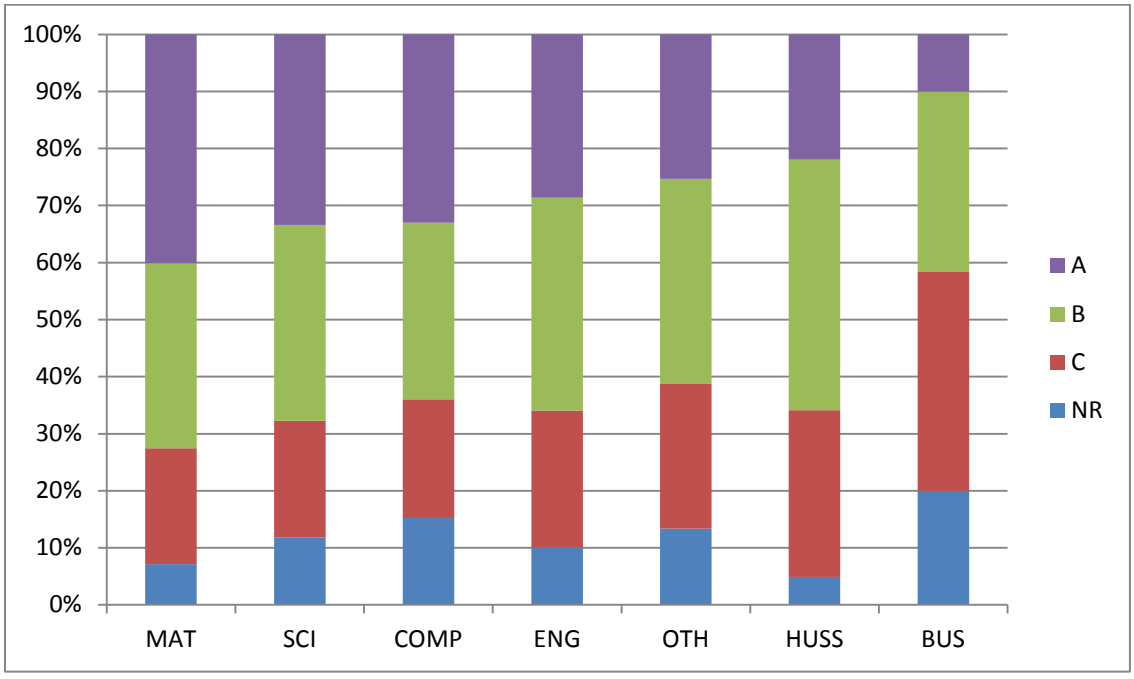


Figure 3: College Major versus Performance

Note: OTH includes undeclared, but not undeclared engineering or undeclared science

Class Year/ Upperclassmen Analysis

The class year testing was very interesting because there are so many different variables involved. I have found that the longer I have been at Worcester Polytechnic Institute the more successful I am, even though the classes have gotten much harder. In the class year data analysis graph, Figure 4, it shows that freshmen are more successful than upperclassmen. That being said freshmen have a higher percentage of A's and B's and lower percentages of C's and NR's. There are multiple reasons for why the data may look this way. First of all the basic level physics courses are generally taken during your freshman year and are much easier than the upper level physics courses which students would take as upperclassmen. Also upperclassmen taking basic physics courses are most likely students who failed the class freshmen year or changed majors because they were not doing well in school. In either case these students are likely to bring down the class averages just a bit. The most significant result from this data in my opinion is the upperclassmen fail physics classes at about double the rate of freshmen. I was very surprised to discover these findings; I know the physics classes are generally more basic freshmen year however I personally have had much more success as an upperclassmen. Although I am probably an odd exception, as the data shows, I know many upperclassmen that slack off and care about grades much less than they did freshmen year. An interesting experiment would be to look into students who took Physics 1110 freshmen year and students who took Physics 1110 as an upperclassmen, never having taken it before, and compare those results. They may be slightly different than the ones in Table 1 and Figure 1 simply because no harder courses are involved and no students who failed the class are involved in the study. The appendix shows a histogram of first year students taking upper level physics courses versus upperclassmen taking upper level physics courses. The results show the first year students excelling, this shows that Worcester

Polytechnic Institute's physics placement program is doing a wonderful job helping students be placed in the correct physics course coming out of high school and into college.

Count of SHVGPAI_GRADE	**Column Labels**				
Row Labels	**NR**	**C**	**B**	**A**	**Grand Total**
F	631	1582	2654	2247	7114
U	656	906	1266	1136	3964
Grand Total	**1287**	**2488**	**3920**	**3383**	**11078**

Table 4: Freshmen versus Upperclassmen

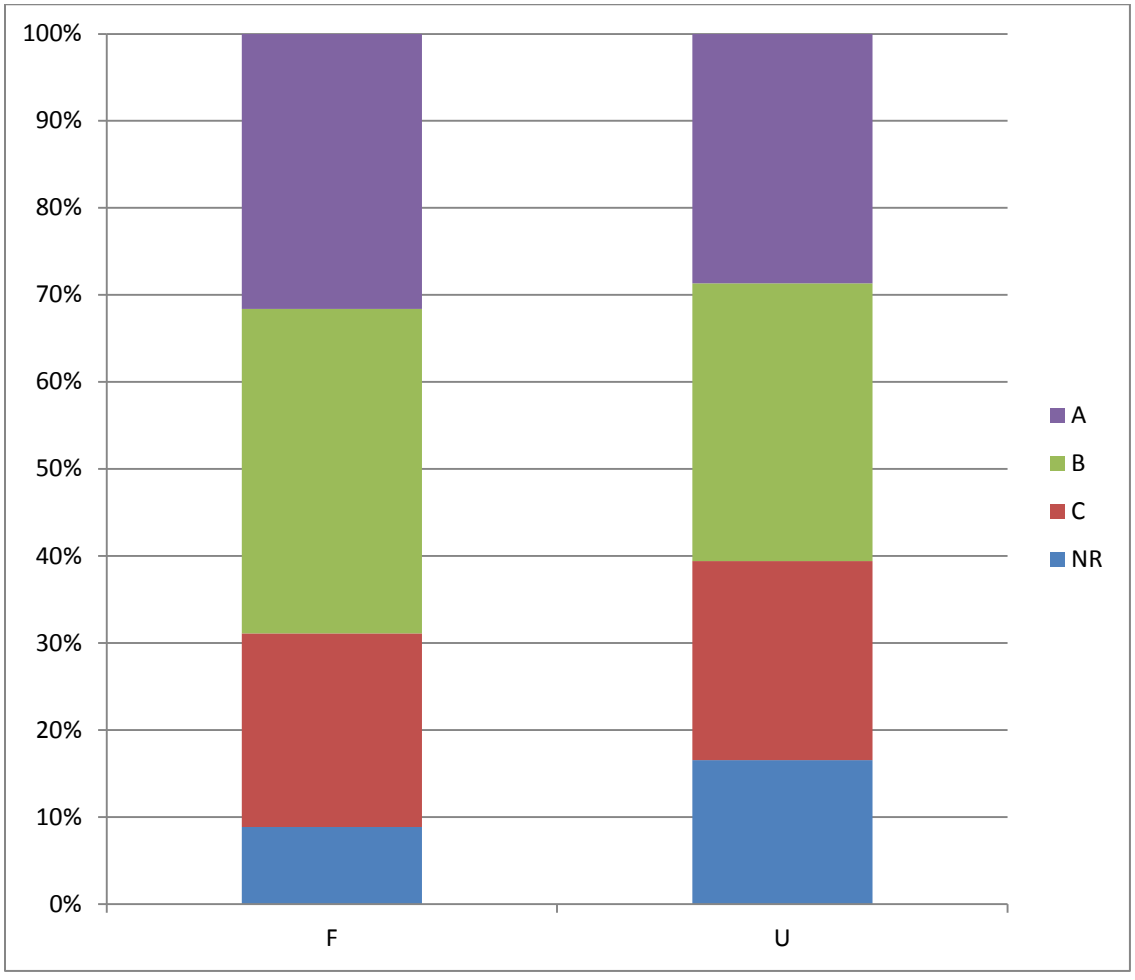


Figure 4: Freshmen performance versus Upperclassmen Performance

Calculus vs. Physics

Calculus is a branch of mathematics that focuses with derivatives and integrals. Calculus is a basic math course that is offered here at Worcester Polytechnic Institute that is usually taken concurrently with basic physics courses. Calculus and Physics concepts coincide with one another and may help with the success of the other course being taken. The calculus versus physics grade graph, Figure 5, is very interesting although the results were not very surprising. The data convincingly shows a direct correlation between physics grades and calculus grades. Eighty one students received an A in both their physics class and their math class. This number was so much larger than any other combination that it must show some correlation between the two. Physics is a very math based science which why I believe the grades match up so well. Also if you are good in math it helps you in physics and if you are good in physics it helps you in many upper level mathematical engineering based courses. Table 5 shows the calculus grade received versus the physics grade the student received. Table 6 shows the calculus grade on a scale from 1 through 4, 1 being a NR, 2 being a C, 3 being a B, and 4 being an A. The other column is the average physics grade received, on a scale from 1 through 4, in relation to their calculus grade. As you can see from Table 6 the grades match up very similarly. To further this study one might look into higher level calculus courses versus higher level physics courses. Also the students major is an unaccounted for variable that may affect the results. A mathematics major would most likely do very well in any calculus course but may struggle with physics even though they are closely related. The appendix graphs for this section are a little harder to interpret. It seems to show however that lower level physics courses have a grade trend that has a

higher tendency to get a C in the course while the upper level physics graph follows more closely to that of the overall trend. This would be something very interesting to look into deeper.

Count of Grade2	**Physics Grade**				
Calc Grade	**A**	**B**	**C**	**NR**	**Grand Total**
A	81	22	9	4	116
B	23	22	19	4	68
C	1	11	12	8	32
NR	2	4	8	8	22
Grand Total	**107**	**59**	**48**	**24**	**238**

Table 5: Calculus Grades versus Physics Grades

	Physics Grade	Average Calc Grade	
NR	1	2.166666667	
C	2	2.604166667	
B	3	3.050847458	
A	4	3.710280374	

Table 6: The Relationship Between Calculus and Physics

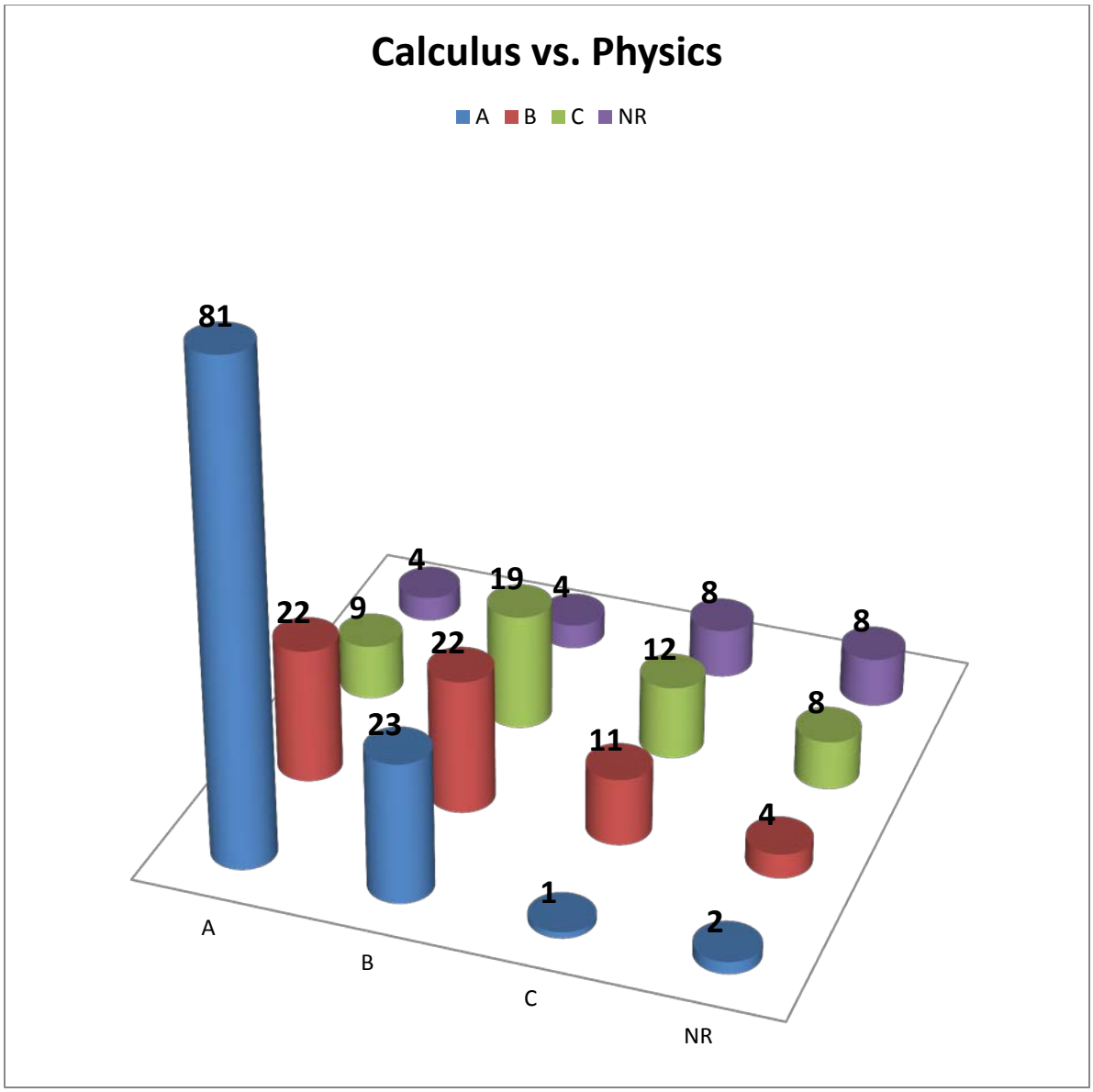


Figure 5: Physics Performance versus Calculus Performance

Conclusions and Recommendations

After playing around with thousands upon thousands of data points I still feel as though there is so much more testing that can be done. The data results that I found were very interesting although not many facts were able to come of the project. I once agreed with the old saying that males are smarter than females and after doing this project I do not agree with it at all. I firmly believe in the data showing the good math scores correlate with good physics scores because both subjects are closely related and that's what I expected the results to look like. Also I expected students with majors more closely related to physics and math to do better than those of majors completely unrelated to physics, and this was true. I did expect upperclassmen to do better than freshman simply because after a year of attending Worcester Polytechnic Institute it is much easier to balance classes and workloads and discover the most successful style of learning for you. However it made perfect sense that freshman performed better because they take much easier physics based classes and there are some students who tend to try less in classes as they progress through college. The appendix shows many more ways the data can be broken down to further dive into explanations for certain statistical trends. There are many more tests I wish I could have run and much more data I wish I had looked into but working by myself and working under time constraints this was the tests that I felt comfortable publishing the results of. I hope that this project is a stepping stone into much more detailed testing of data that can prove the hypotheses that I came up with. I strongly recommend having some previous knowledge with using excel as I did not which made the project a lot harder than expected. I believe my results and the data given to me can help to improve the quality of teaching here at Worcester Polytechnic Institute. Professors who look into the connections in the data would be better able to understand how the learning process works and may be able to change their techniques of

teaching to help create a fair learning environment for all. An example of how to do this may to have a physics class that is much more hands on type of learning so that students who are not as good in math or have a major in a non-related field may be better equipped to succeed. Throughout this project I was continually amazed be the results of the data and I think anyone who ever has a chance to do what I did would love the opportunity.

References

Bryner, Jeanna. Live Science. *Men Smarter than Women, Scientist Claims.*Web. 20 April. 2012. http://www.livescience.com/7154-men-smarter-women-scientist-claims.html

Feuer, L. S., 1974, Einstein and the Generations of Science: New York, Basic Books.

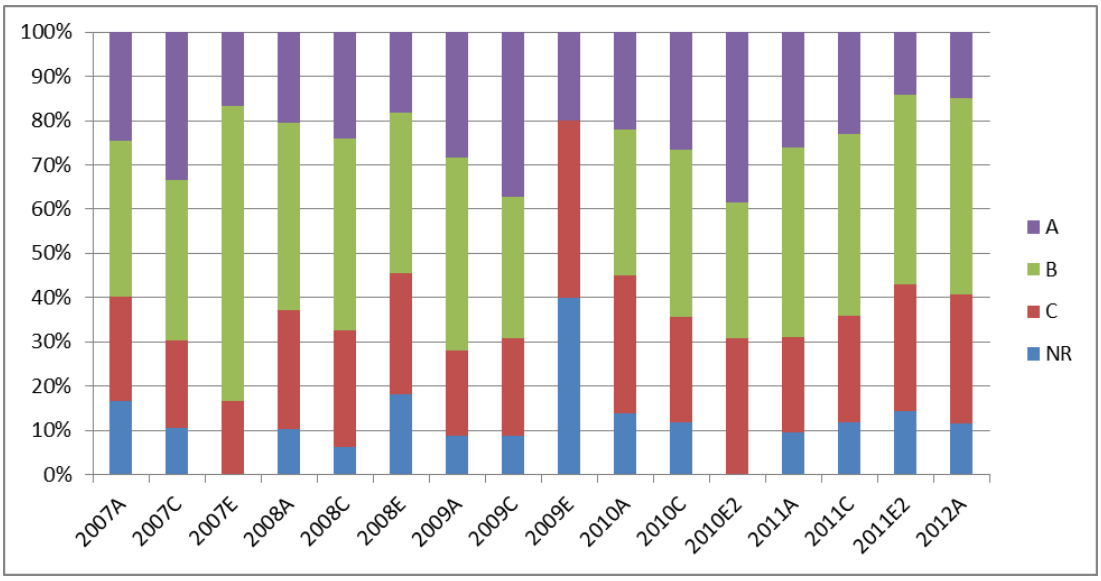
Freedman, A. Roger. Young, D. Hugh. University Physics 13th Edition. Published: Addison-Wesley copyright 2012.

Schlipp, P., 1951, Albert Einstein: Philosopher Scientist: New York, Tudor.

Weidner, R. T., and Sells, R. L., 1975, Elementary Physics: Boston, Mass., Allyn and Bacon.

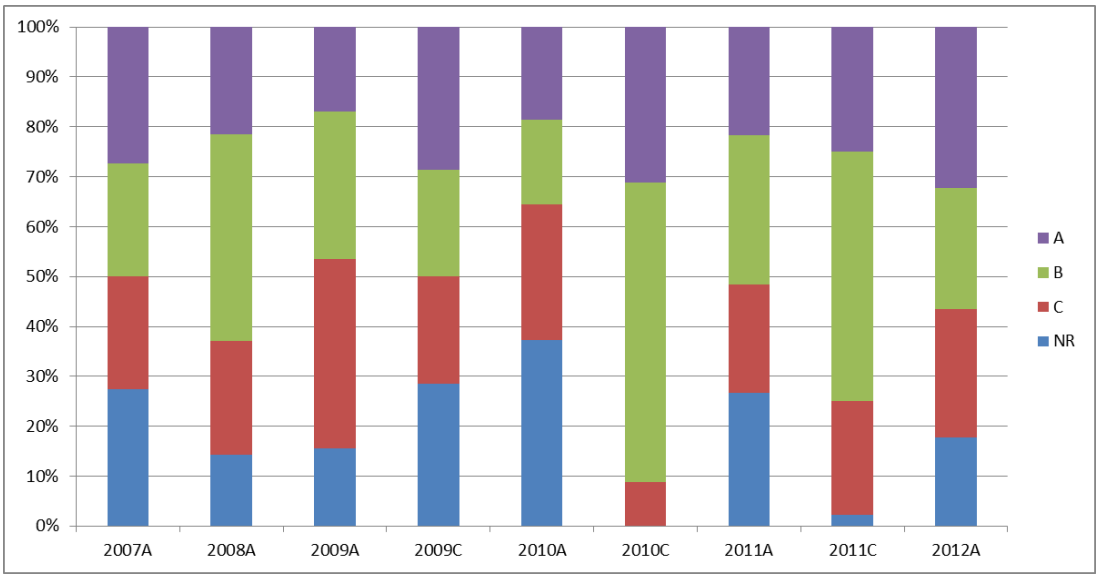
Appendix

Grade Trend

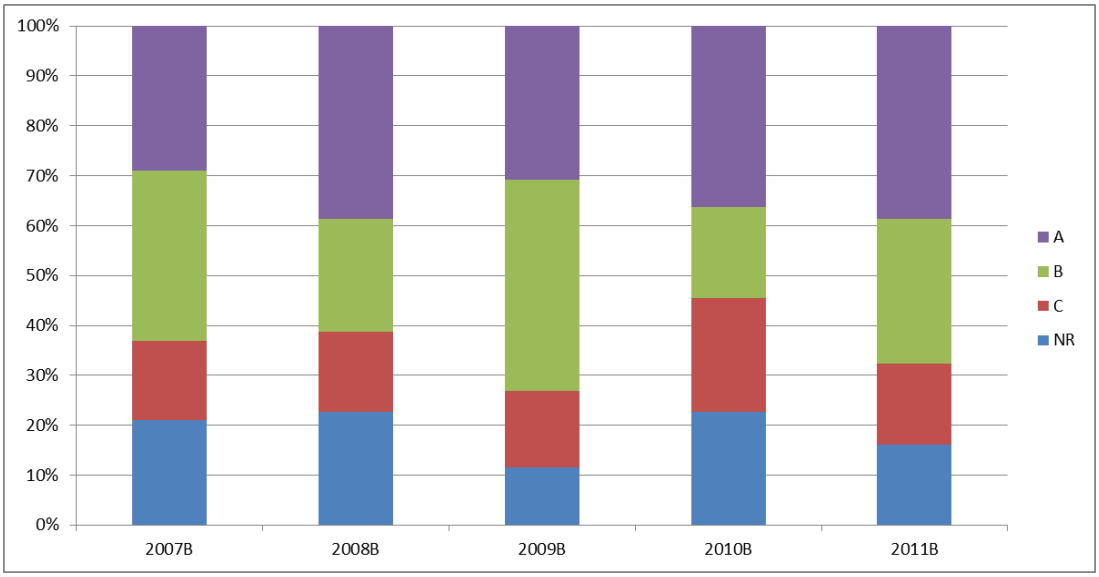


PH 1110

PH 1111

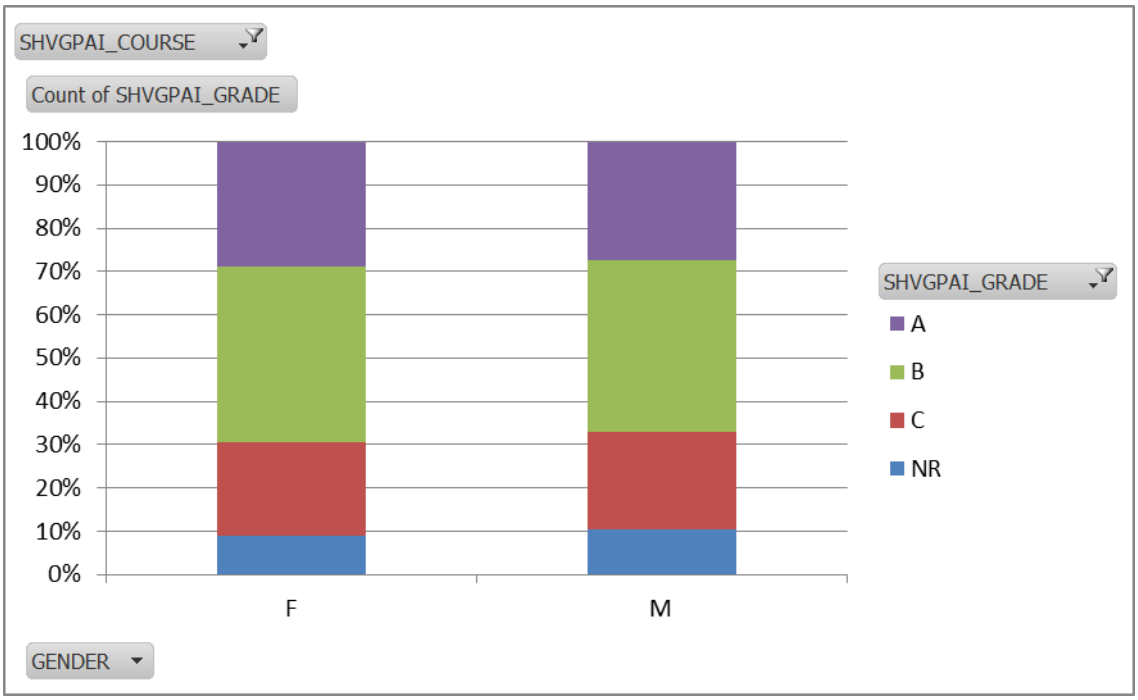


PH 2201

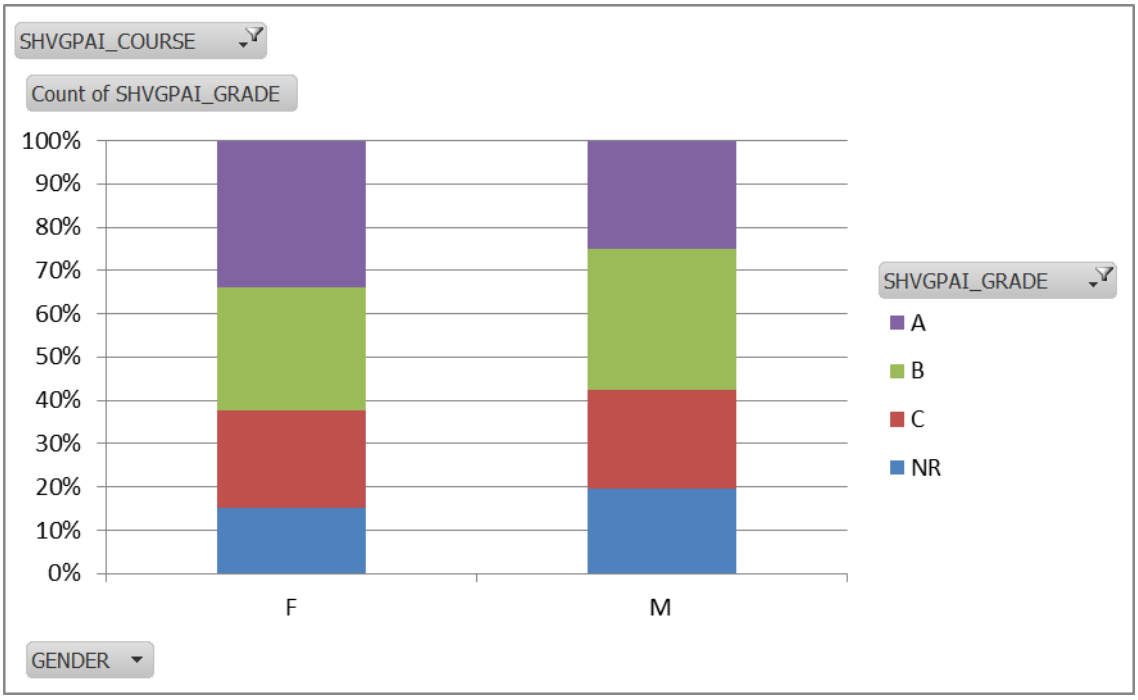


PH 2202

Gender

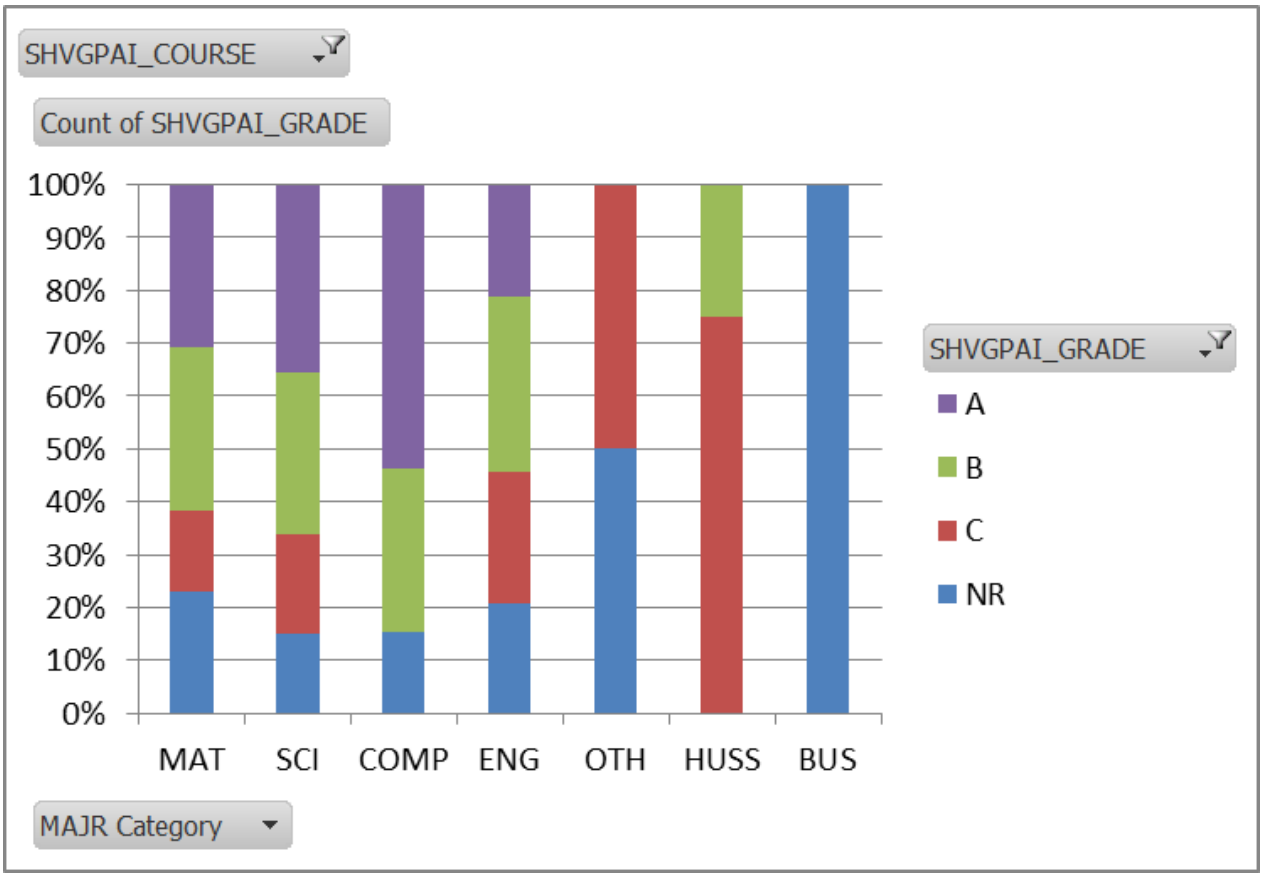


PH 1110 and PH 1111



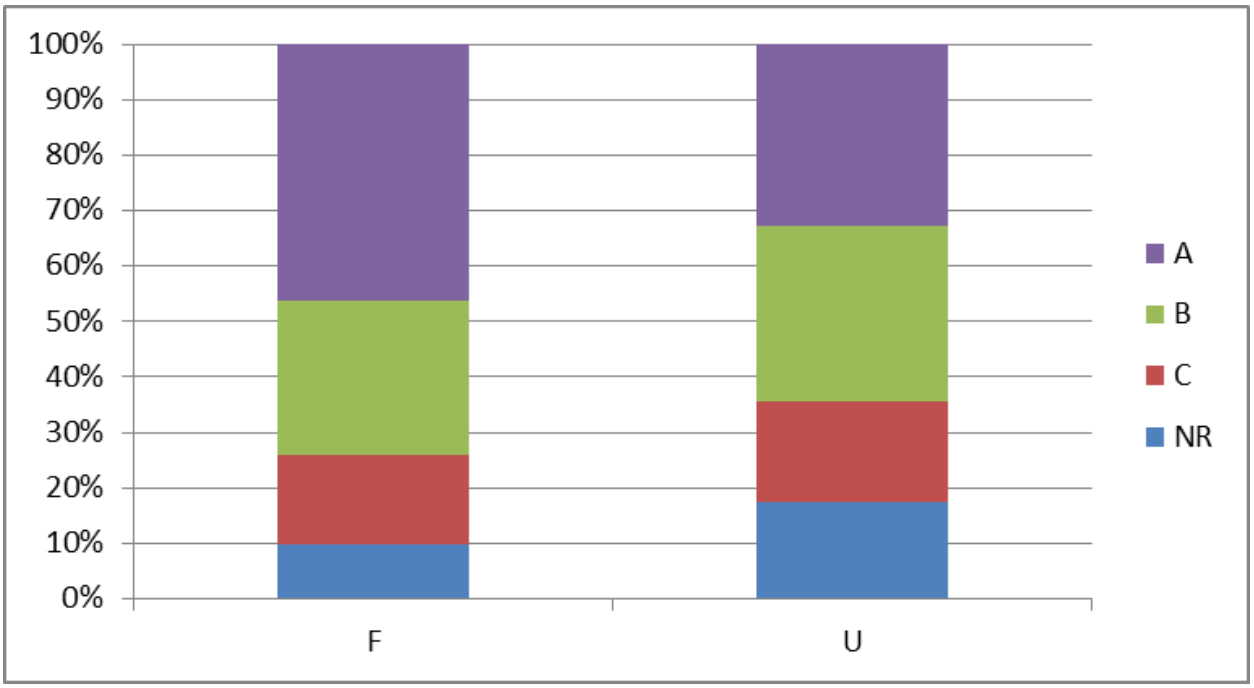
PH 2201 and PH 2202

Major



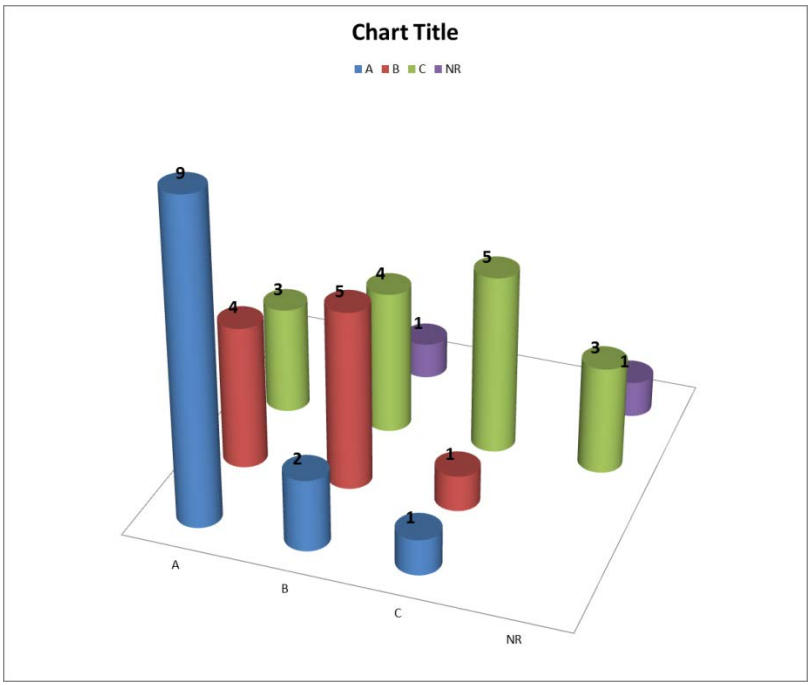
PH 2201 and PH 2202

Class Year/ Upperclassmen Analysis

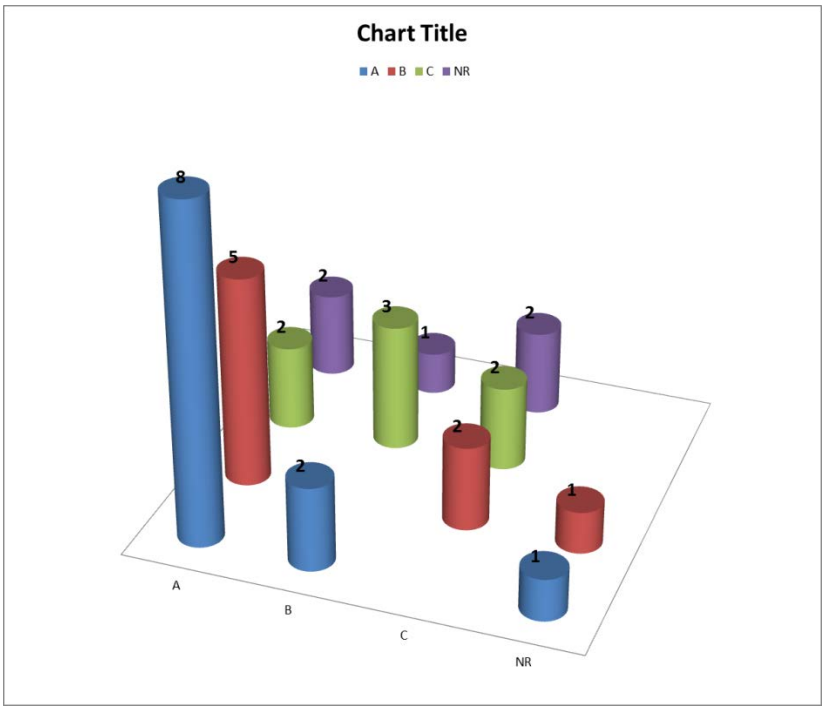


PH 2201 and PH 2202

Calculus vs. Physics



PH 1110 and PH 1111



PH 2201 and PH 2202